DOCUMENT RESUME

ED 280 140 EA 019 243

AUTHOR Wightman, Margaret
TITLE The System-Wide Evaluation Project in Ottawa Board of Education High Schools. Research Report 86-04.
INSTITUTION Ottawa Board of Education (Ontario). Research Centre.
PUB DATE Dec 86
NOTE 52p.; Paper presented at the Annual Meeting of the Ontario Educational Research Council (December 1986).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Academic Achievement; *Cooperative Planning; Educational Improvement; *Evaluation Methods; Foreign Countries; High Schools; Professional Development; *Program Development; Program Effectiveness; Program Implementation; *State Programs; Teacher Participation; *Test Construction; Test Interpretation; Test Results
IDENTIFIERS *Ontario

ABSTRACT

This report describes the procedures of a systemwide testing program established by the Ottawa, Ontario, education board to evaluate high school student performance in selected courses. The project was initiated in response to a directive from trustees, who were concerned about variation of academic standards among students in relation to university admission criteria. The report analyzes program procedures and presents a sample of results obtained and some reactions of project participants. Tables and figures present student evaluation data throughout the report. The study is organized into the following sections: (1) schedule of testing; (2) test construction, including development, school preparation, and scoring; (3) test content; (4) processing data, including actual and final scores, test-specific results, and rescoring; (5) professional development activities, including the test construction committee, teacher workshops, posttest meetings, and inschool meetings; and (6) views of participants, including those of administrators, consultants, principals, teachers, and students. The study concludes that the high degree of project flexibility permitted the diverse committees to work independently and to choose appropriate procedures. This approach thereby promoted the program's goal of excellence in evaluation, and followup activities continue. Teacher involvement, essential to the project's orientation, should assist in elevating academic standards. Outlines of tests by course area are appended. (CJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED280140

THE SYSTEM-WIDE EVALUATION PROJECT

IN OTTAWA BOARD OF EDUCATION HIGH SCHOOLS

by

Margaret Wightman

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Margaret Wightman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

RESEARCH REPORT 86-04

EA 019 243

A Paper Presented at the Annual Conference
of the Ontario Educational Research Council, December 1986.

research centre / centre de recherches
The Ottawa Board of Education / Le Conseil scolaire d'Ottawa
Location: 662 Lyon St. S., K1S 3Z2, tel. 563-2408 Mailing address: 330 Gilmour St., K2P 0P9

BEST COPY AVAILABLE

TABLE OF CONTENTS

	Page
INTRODUCTION	1
A Brief Description	1
The Background	1
SCHEDULE OF TESTING	2
TEST CONSTRUCTION	6
The Committee	6
Common Course Outline	6
Test Development	8
Sources of Items	9
Translation	10
School Preparation	10
Scoring	11
TEST CONTENT	12
PROCESSING THE DATA	13
Actual and Final Scores	13
Estimated Score	14
School Measures	15
Test Specific Results	22
Rescoring	23
PROFESSIONAL DEVELOPMENT ACTIVITIES	27
Fall Conference	27
Tests Construction Committee	27
Teacher Workshops	28
Pooled Marking in English	29
Post-test Meetings	30
In-school Meetings	31
VIEWS OF PARTICIPANTS	32
The Administration	32
Consultants	34
Principals	34
Teachers	35
Students	37
CONCLUSION	38
APPENDIX	40

LIST OF TABLES

		Page
Table 1	SYSTEM-WIDE STUDENT EVALUATION IN THE OTTAWA BOARD OF EDUCATION - SCHEDULE OF TESTING FOR ADVANCED LEVEL COURSES	4
Table 2	SYSTEM-WIDE STUDENT EVALUATION IN THE OTTAWA BOARD OF EDUCATION - SCHEDULE OF TESTING FOR GENERAL LEVEL COURSES	5
Table 3	OVERVIEW OF DIFFERENCES IN MALE AND FEMALE MEAN SCORES - SYSTEM-WIDE STUDENT EVALUATION RESULTS, 1985	21
Table 4	CORRELATION COEFFICIENTS FOR FRANCAIS 12A - SYSTEM-WIDE STUDENT EVALUATION RESULTS, 1985. .	22
Table 5	RESULTS OF STUDENT QUESTIONNAIRE IN PHYSICS 11A - SWSE TEST OF JUNE 1986	24
Table 6	RESULTS BY TOPIC AND SEX FOR MULTIPLE-CHOICE ITEMS IN PHYSICS 11A - SWSE TEST OF JUNE 1986	25
Table 7	RESULTS BY SEX FOR PHYSICS 11A - SWSE TEST OF JUNE 1986.	25

LIST OF FIGURES

		Page
Figure 1	RESULTS OF SYSTEM-WIDE STUDENT EVALUATION - PROFILE OF ALL SCHOOLS, JANUARY 1986	16
Figure 2	RESULTS OF SYSTEM-WIDE STUDENT EVALUATION - PROFILE OF ONE SCHOOL, JUNE 1986	17
Figure 3	SYSTEM-WIDE STUDENT EVALUATION - RANGE OF SCHOOL AVERAGES, JUNE 1986	18
Figure 4	PLOT OF BOARD AND SCHOOL RESULTS ON GEOGRAPHY 9A, JANUARY 1986, SHOWING MEDIAN AND 90th, 75th, 25th AND 10th PERCENTILE LEVELS	20

INTRODUCTION

A Brief Description. In the fall of 1984, the Ottawa Board of Education (OBE) set up a board-wide program of testing in the high schools that has become known as the System-Wide Student Evaluation (SWSE) project. The original purpose of the project was to have some common tests administered to all students in OBE high schools who were taking certain selected courses in core subjects. A fairly small number of courses was targetted for testing during the first three years but these were drawn from all grades and both advanced and general level courses. The tests are scheduled for sometime between January and June and the mark received on the SWSE test becomes a set portion of the students' final mark in the course.

The Background. This program of system-wide evaluation was developed in response to a directive from the trustees who were concerned about a possible variation of academic standards amongst OBE high schools. In the months previous to the establishment of a testing program the press had carried several reports from universities about the lack of uniformity on the part of high schools in awarding grade 13 marks. Suggestions were made that, when admitting students, universities were placing considerable importance on the particular high school attended as well as on the marks obtained there. At least one university reported the development of a formal rating scale of schools which was presumed to be used to adjust the marks submitted in entrance applications. Trustees felt that parents needed to be reassured that their children had the opportunity to obtain a good secondary school educa-

tion, no matter which OBE school they attended. Similar concerns in other Canadian educational jurisdictions had recently led to broad-based testing so it was not too surprising to see such a program established in the OBE. The Superintendent of Programme become responsible for the SWSE and one of the three administrative assistants in the Programme Department coordinates the project.

SCHEDULE OF TESTING

The project administrators, after consultation with principals, subject consultants and others, selected the courses to be tested for the first three years of the program. These are shown in Table 1 for advanced level courses and in Table 2 for general level courses. As well as identifying the course, the tables show the year and month the test is given, the time allowed for the test and the percentage of the final mark in that subject that will be determined by the SWSE test score.

It can be seen in Tables 1 and 2 that the first year differed from later years in that there were several periods of testing between January and the end of May. This pattern was chosen, in part, to avoid overloading the various support systems that would be required for the testing program. Also, having tests before the end of the year allowed for the possibility of some to be used as a formative evaluation tool with results being available to teachers while they were still instructing the tested students.

After the first year, the SWSE tests have been scheduled for two distinct periods, one in January and another in June. This allows the SWSE tests to be incorporated into the regular midyear examination

schedule if these exams are held in January. At present some OBE high schools have midyear exams in December and others in January. Except for unusual circumstances, all schools must participate in the SWSE testing and it is a policy that a student will not be asked to write another major test for the same course within a one month period of the SWSE test. For these reasons the SWSE project may promote a common midyear examination period across the board if principals and teachers decide they would like the SWSE tests and the school midyear examinations to be held at the same time. However, in year three of the project, only a limited movement in this direction has taken place.

Another difference in the first year of testing, shown in Tables 1 and 2, is that two tests were given for each of English 9A and 9G, Français 9A and 9G and Mathematics 9G. It was felt that having two sittings instead of one should reduce the students' anxiety and consequently enhance their performance. It was expected that some of the grade 9 students would not have previously experienced a similar testing situation. In the case of the grade 9 test of English for G level students, the second test had the same format but different content. It took place one week after the first test and the teacher reported the better test score. Having two test periods was appropriate for some other SWSE tests in the language arts because the first was used to prepare drafts and the second, to complete the writing assignments.

TABLE 1

SYSTEM-WIDE STUDENT EVALUATION IN THE OTTAWA BOARD OF EDUCATION
SCHEDULE OF TESTING FOR ADVANCED LEVEL COURSES

Subject	Grade	Time of testing			Length (hrs)	Weight (%)
		1985	1986	1987		
English	9	Apr.	-	-	1.5	15
	9	May	-	-	1.0	10
	10	-	-	Jan.	1.5	15
	11	-	June	-	2.0	20
	12	Mar.	-	-	2.0	25
	12	-	June	-	2.0	30
	12	-	-	June	2.0	30
Français	9	Jan.	-	-	1.0	10
	9	May	-	-	1.5	15
	10	-	-	Jan.	1.5	15
	11	-	Jan.	-	2.0	25
	12	Mar.	-	-	2.0	25
	12	-	June	-	2.0	30
	12	-	-	June	2.0	30
Mathematics	9	Apr.	-	-	2.0	25
	12	Mar.	-	-	2.0	25
	12	-	June	-	2.5	30
	12	-	-	June	2.0	30
Calculus	13	-	June	-	2.5	35
Science	9	Apr.	-	-	2.0	25
	10	-	-	Jan.	1.5	15
Physics	11	-	June	-	2.0	30
Chemistry	12	Mar.	-	-	2.0	25
Physics	13	-	-	June	2.5	30
Geography	9	-	Jan.	-	1.5	15
History	10	-	-	Jan.	1.5	15

TABLE 2

SYSTEM-WIDE STUDENT EVALUATION IN THE OTTAWA BOARD OF EDUCATION
SCHEDULE OF TESTING FOR GENERAL LEVEL COURSES

Subject	Grade	Time of Testing			Length (hrs)	Weight (%)
		1985	1986	1987		
English	9	Apr.	-	-	1.0	10
	9	May	-	-	1.0	10
Français	9	Jan.	-	-	1.0	10
	9	May	-	-	1.0	10
Mathematics	9	Jan.	-	-	1.0	10
	9	Apr.	-	-	1.0	10
Science	10	-	-	Jan.	1.0	10
Geography	9	-	Jan.	-	1.0	10
History	10	-	-	Jan.	1.0	10

TEST CONSTRUCTION

The Committee. The tests used in the SWSE project are contructed by members of a Test Construction Committee. The chairperson is a vice-principal who has taught in the same subject area as the test to be constructed. The subject consultant is also on the committee and carries most of the responsibility for preparing and distributing the materials used by the committee. The other members of the committee are experienced and respected teachers. There are usually three teachers appointed for each test. In the case of testing A and G level students in the same grade and subject, often a joint committee is struck with one chairperson and consultant but a double complement of teachers, half representing each level. In subjects other than English and Français (which is the French language arts course for Francophone students) at least one of the teachers on the committee is from a French high school.

A resource person from the Research Department is also available to each committee but the involvement of this person has varied greatly from one committee to another. On some committees the member from Research has participated fully in the test construction phase while on others there has been little or no contact with Research personnel.

Common Course Outline. The first task of the Test Construction Committee is to establish a common course outline based on the curriculum guidelines produced by the Ontario Ministry of Education. The high schools in the Ottawa board have been required to develop, independently, an outline for each course offered and sometimes these documents have been used by the committees to prepare a common course

outline to be used in all schools. In other cases much of the document was already prepared, sometimes by the consultant, before the committee met. Once a common course outline is accepted by the committee it is sent to the schools well in advance of the SWSE test. All teachers are informed that this outline should be used to determine course content and that it will also serve as the guide for constructing the test. As part of the common course of study many committees also developed or accepted a set of testable objectives for the course. The source of these objectives varied but when they were used by committees they were also sent to other teachers of the course.

This process of developing a course outline, as with many others in the SWSE project, varied greatly in difficulty and time requirements from one committee to another. One group which found drawing up the outline to be particularly difficult was the grade 10 History Test Committee. The course, titled Contemporary Canada, deals with four major themes in contemporary Canadian history; relations between English and French Canadians, Canadian government and law, relations with our American neighbours, and Canada's role in world affairs. Needless to say, a course that includes the relationships between different peoples is very apt to have a variety of presentations and in a board with both French and English high schools this diversity was accentuated. As well as the different cultural backgrounds of the two groups of teachers, each approach the course with quite different resources. Very few French texts are available whereas there are many English. It was the observation of the consultant in this area that the difference in the amount and type of material available to the two language groups may be responsible for the contrasting patterns in course organization

which tend be chronological in the English domain and thematic in the French.

Test development. The format and content of the test used for each subject is determined by the Test Construction Committee. The common course outline, together with any course objectives, serve as a guide to topics that may be tested. It is not unusual for a committee to find that an item that seems very appropriate for a test is not included in the descriptions of the course. Such an item is not used.

In drawing up a test the committees have many factors to take into account. The topics tested need to reflect the course content in scope and balance. This is sometimes achieved by matching the proportion of class time devoted to various sections of the course to the proportion of the total score determined by items dealing with the same topic. Normally, not every topic in a course is tested but topics are selected from every major section of the course. The style of individual items also must be given careful consideration. Efforts are made to test a variety of learning skills and to avoid the overuse of items aimed primarily at testing knowledge. The overall test design must be appropriate for the subject and students as well as being easy to administer. Field-testing an instrument before its use in the SWSE context has taken place very infrequently. One thing that guides much of the work of the committee is the anticipated reaction of other colleagues teaching in the same field. To a certain extent, the committee members aim to produce a test which will serve as an exemplary model of its kind.

Sources of items. The items used on the tests come from a variety of scources. It is the teachers on the Test Construction Committee, sometimes helped by the consultant, who bring forward draft items. Often teachers divide this search for items by each agreeing to cover certain teaching units or sections of the test. They use any appropriate resources available to them including their own skill in developing new items. Some items are taken as previously used or as developed but many more are adapted and "fine-tuned" for the particular test by the committee.

One source of draft items has been the Ontario Assessment Instrument Pool (OAIP). This resource was used extensively for the Geography, Physics and Chemistry tests. In these subjects the greatest value in having OAIP was felt to be in providing multiple-choice items. This form of item is known to be easier to score but usually much more difficult to construct than other types so teachers can save considerable amounts of time when a large pool of such items is available. Other styles of OAIP items seem to serve primarily as a model or inspiration in developing new items. OAIP packages are not available for some of the courses tested by the project.

A source of items for one of the third year Mathematics tests will be the collection developped by teachers attending a half-day workshop. About 50 teachers were divided into groups of 6 or 7 to produce items based on the course objectives for a particular section of a course. This event will be described more fully in the discussion of professional development activities.

Translation. Both a French and English version of the test are required for all courses except the language arts. In general teachers on the test construction committees develop and submit items in their first language although the working language of most committees is English. Some committees have produced parallel tests in both languages but the more common occurance is for a test to be produced in English and then translated into French. Although the translated version is usually scrutenized very carefully by several persons whose first language is French and who also know the subject area, there have been complaints voiced about some details of translations. For example, in the Science 9A test it was felt by certain teachers that the level of vocabulary used was too difficult for some of the local French-speaking students. Many of these students are bilingual and use French less frequently than they would if they lived in a predominantly French-speaking part of the country.

School Preparation. As mentioned previously, the teacher of a course to be tested on the SWSE project receives a common course outline describing the scope of the course as early as possible. When the test is drawn up, a description of its general nature is usually sent to the schools. This gives an outline of the test format with a breakdown of the marks that would be awarded for each section. Other information may be included such as the emphasis or ommission of certain sections of the curriculum or a preview of the instructions for marking that will accompany the test. In some cases a practise test has been distributed to schools. This is made up of items that have been developed but not used on the actual SWSE test. In general, the committees try

to reduce the apprehension felt by teachers and students by giving them a reasonable amount of information about what to expect on the test.

Scoring. There are facilities in the board to have standard answer sheets scored by an optical scanner and the results of this process printed on copy that can be returned to teachers. No supplementary resources are available for external marking for SWSE so that, other than multiple-choice items, the tests must be scored by the teachers themselves.

When the test is finished, the committee prepares a marking scheme. For short answer questions, a copy of correctly answered test items is often used. Sometimes there is additional information about the requirements for each point awarded. The correct answers for large numbers of multiple-choice items are sometimes accompanied by a template so that teachers can score this section of the test, if they wish, before sending the answer sheets to the Programme Department for optical scanning. For essay-type questions there is usually a breakdown of the marks to be awarded with a description of what is expected in each category.

Some committees have set up a marking workshop either before or immediately following the writing of the test. These allow teachers to discuss any issues of the marking that need clarification. The other variation to the normal marking procedure is to have a pooled system of marking with teachers scoring the tests of students not in their classes. This was actively encouraged during the first year and its use for part of the English 12A test will be described in a later section of the report.

TEST CONTENT

The tests that have been produced by the Test Construction Committees have varied widely, especially from one subject area to another. An outline of the contents of each test given in the first two years of the SWSE Project is given in Appendix A. Table A-1 deals with the advanced level courses and Table A-2 with the general level courses. For each test a brief description of the major sections is given along with some information about the evaluation procedure. Also given is the total value of the section of the test and the suggested time allowance for the section when this was indicated on the test.

It can be seen in Appendix A that the tests include a variety of item types but there are also some common features, particularly within certain subject areas. The tests of English and Français each include at least one sight reading of prose or poetry. To determine the comprehension of these pieces, multiple-choice items are often used for English and short-answer items for Français. Short essays may also be used to test the sight-read piece. A longer essay or composition also appears on every test of language arts. In English, one essay is usually based on a Shakespearian play or a novel studied during the course. In Français for grades 9 and 11, there is a section of the test which measures grammatical knowledge in such areas as homonyms (i.e. peu, peux, peut), verb forms and agreement of adjectives. Both tests of Français for 9G students and one of the two tests for 9A students include a section based on a recording which gives instructions to students for completing a response sheet. Students were not allowed to have dictionaries for tests in the language arts.

Most of the tests in Mathematics have three sections; the first is multiple-choice, the second is short answer and the third is calculations or problems that require complete solutions . The Calculus 13A test has a second section of short answer items replacing the multiple-choice section of other Mathematics A level tests. The grade 9G test in Mathematics has only two sections, one that is mainly filling in blanks and the other calculations and problems requiring complete solutions.

Multiple-choice items make up 50 per cent of the Physics 11A and Chemistry 12A tests and 40 per cent of the Science 9A test. Short answer items make up the rest of Science test and are used for the Physics and Chemistry tests, which also contain problems.

The grade 9 Geography tests are made up entirely of multiple-choice items. There are 100 items for the A level test and 70 for the G level.

During the first two years of the project, the tests for G level students were all given in grade 9. Their overall structure is often similar to the advanced level tests but the content is different. The G level students are usually given shorter pieces of sight reading in the language arts and the tasks are broken down to a greater degree. Usually a smaller amount of writing is required and these tests were written in one hour.

PROCESSING THE DATA

Actual and Final Scores. The 'actual' score is the mark given by the teacher on the SWSE test. Before submitting the set of scores for a class, the teacher temporarily returns the marked test to the students

so that they can see the basis for their mark and discuss any concerns they have with the teacher.

The second score that is associated with each test is the final mark for the course given in June. This data is obtained from the school records stored on the central computer system. Part of the final score is based on the SWSE test; and this percentage can be seen in Tables 1 and 2 on pages 4 and 5.

Estimated Score. During the first year and a half of the project the teachers were asked to produce an estimated mark for each of their students before the SWSE test had been written. The estimated mark was to be based on the students' apparent mastery of the content of the course to be tested. These marks were used to obtain an estimated mean score for all students who would write the test. When the mean score of the actual test marks was available, scaling of marks to the estimated distribution was done if there was a difference of more than five points between the two mean scores. The scaling was done across the board so that the rank order of students would not change but the mean of the scaled scores would be the same as the mean of the estimated scores. This process served as a safeguard against having board-wide results that were much higher or lower than would normally be expected for the course.

During the first year of testing it was necessary to scale the scores for five of the 16 tests. Scores were raised for four tests and lowered for one. As might be expected there was a considerable reaction from schools and parents when the directive was received to lower the marks.

School Measures. The sets of test scores and any other data generated by the SWSE project is transferred to the Research Department for processing. Basic statistics for each school and the board as a whole are produced for each test and tables are generated. Principals receive these results for their own schools and the board as a whole. Consultants receive results for their own subjects. In any other presentation of the school statistics the individual schools are not identified.

Figure 1, for instance, was used to display the distribution of school means for January of the second year of testing when the Superintendent of Programme made his annual report to a meeting of trustees which was open to the public. Each school mean is represented by a dot and the board mean by an arrow. It is possible to see the range of school means and to get an impression of the extent of clustering. It is also possible to see changes in these characteristics from one distribution to another.

Figure 2 contains similar information about the June 1986 school means but, in addition, all the dots representing a mean of a particular school have been circled. It is then possible to see the school's position relative to other schools on the two sets of scores associated with each test. It is also possible to gain an impression of the extent of variation in the school's performance from one test to another.

Another visual representation of SWSE results based on school means is a 'box and whisker' graph. Figure 3 gives such a graph for SWSE tests written in June 1986. The upper and lower edges of the box represent the 75th and 25th percentiles in the distribution of school means and the solid line running across the box represents the median.

FIGURE 1

RESULTS OF SYSTEM-WIDE STUDENT EVALUATION - PROFILE OF ALL SCHOOLS

JANUARY 1986

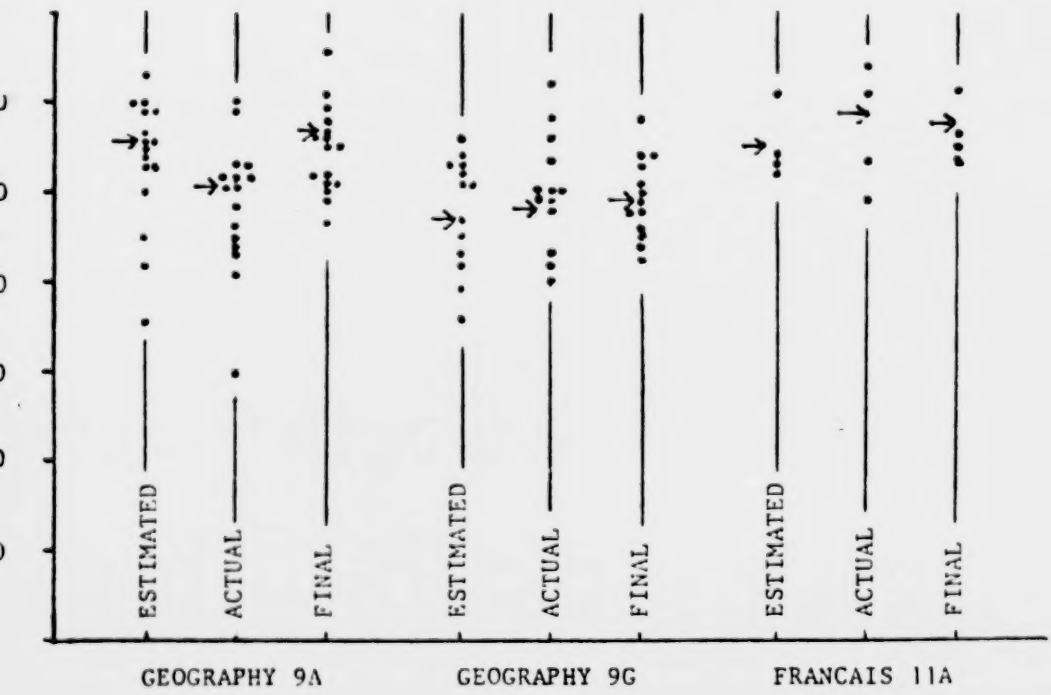


FIGURE 2

RESULTS OF SYSTEM-WIDE STUDENT EVALUATION - PROFILE OF ONE SCHOOL

JUNE 19[illegible]

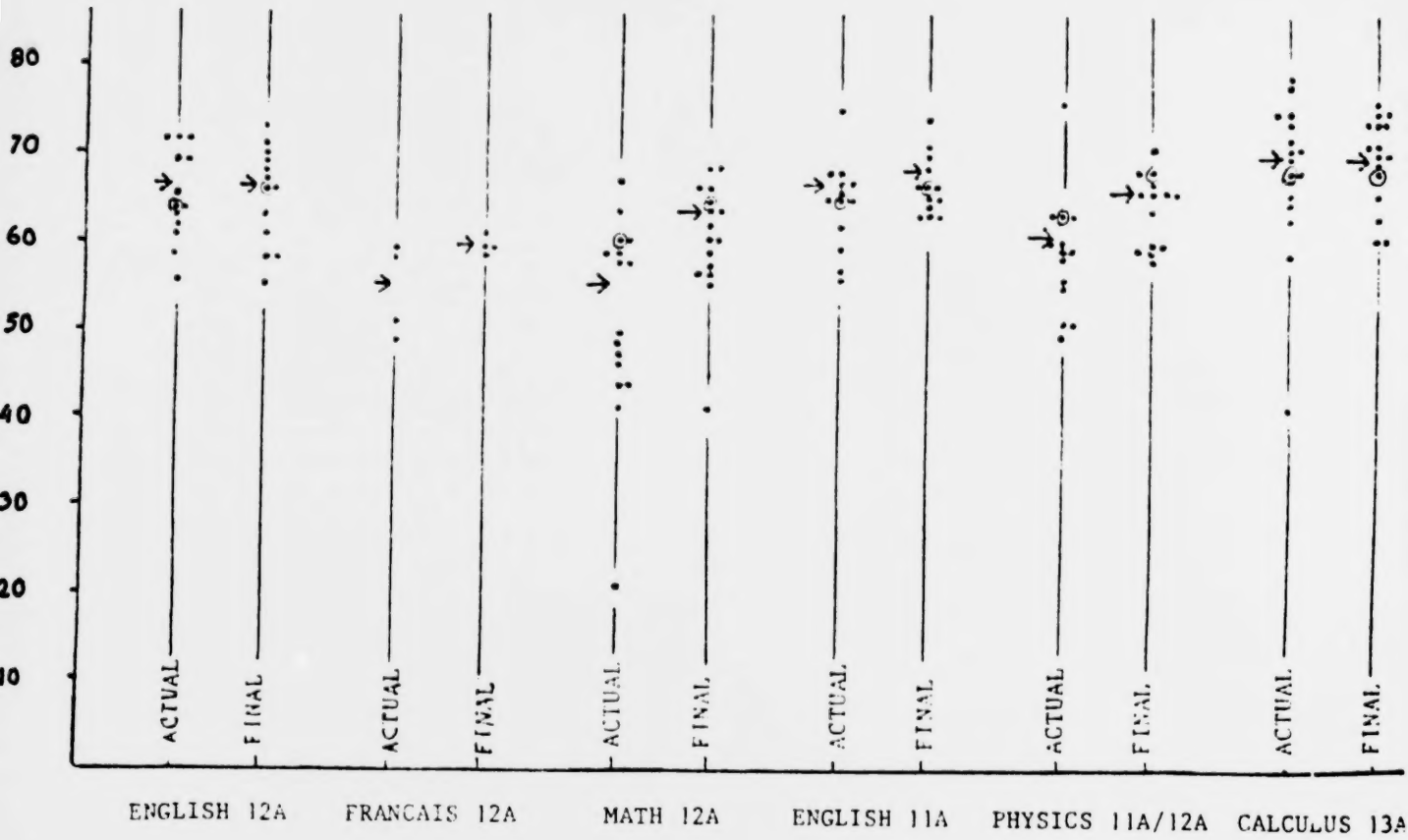
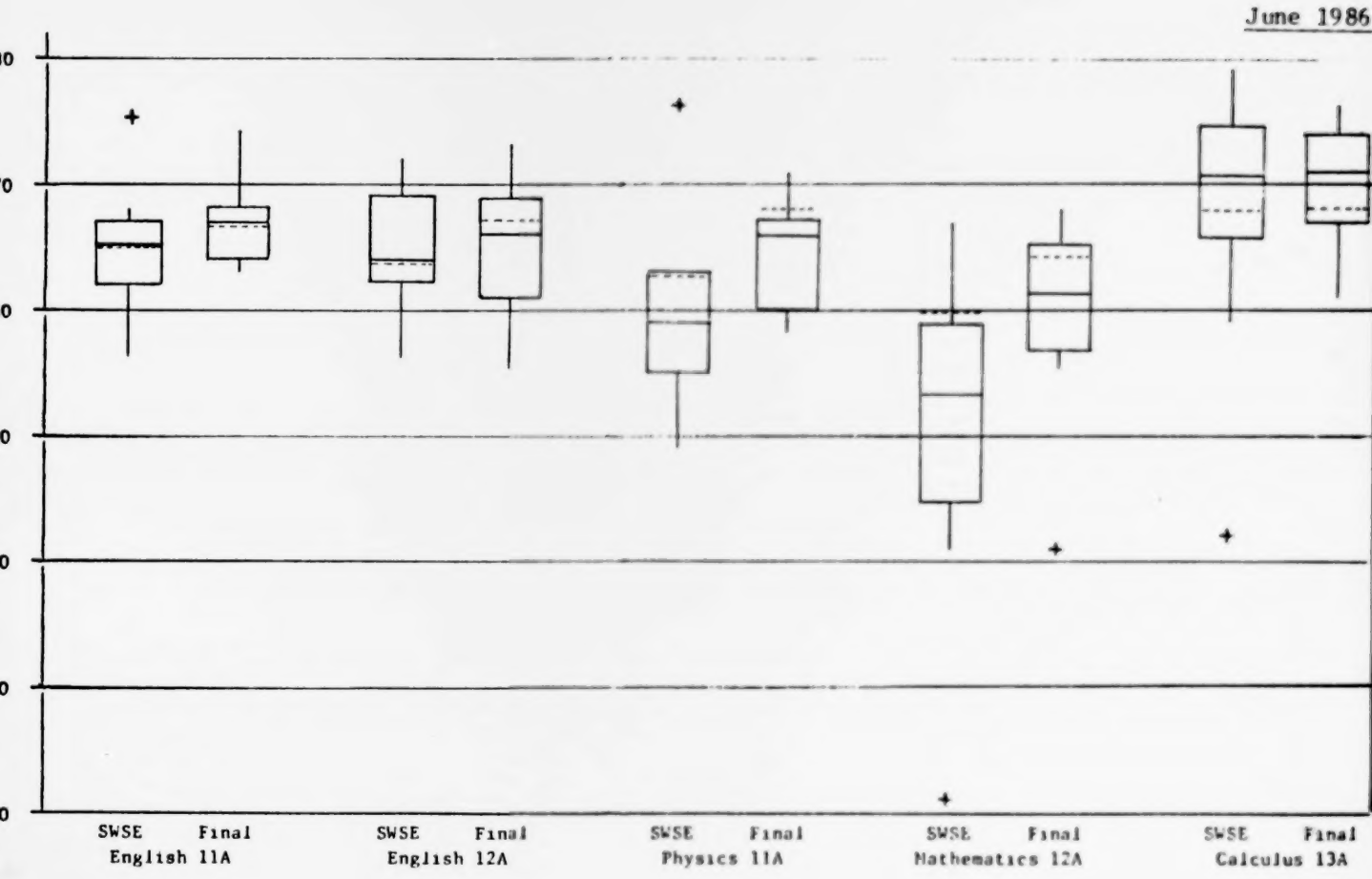


FIGURE 3

SYSTEM-WIDE STUDENT EVALUATION -- RANGE OF SCHOOL AVERAGES



The upper whisker extension represents the highest 25 per cent of school means and the lower whisker, the lowest 25 per cent. The plus sign is used to represent the mean of a school whose score was so different from other means represented in the whisker that its inclusion in the whisker would be a distortion. The dotted lines indicate the means of a particular school and demonstrate how this graph could be used to make comparisons between schools or within a school.

The graph in Figure 4 is similar in appearance but in this case several schools may be compared on the same measure. The graph is based on the results of the Geography 9A test and was prepared for a meeting of department heads. Each figure in the graph represents a school with the range of the 25th to 75th percentile scores shown by the box and the 10th to 90th percentile scores by the whiskers. The horizontal line drawn across the box is the median. This presentation gives a visual impression of the extent to which scores from the various schools overlap.

It is an OBE policy to analyze all test data by sex and an overview of the outcome of this process for the SWSE project in year one is presented in Table 3. Correlation coefficients were calculated at the end of the first year using the actual test results, the estimated scores and the June mark in the course. Table 4 gives the results of this procedure for the Français 12A test. It can be seen that the coefficient is generally lowest between the actual and estimated scores and highest between the estimated and end-of-year scores. A similar observation could be made for all the first year tests although the differences within a set of three coefficients are generally larger

FIGURE 4

PLOT OF BOARD AND SCHOOL RESULTS ON GEOGRAPHY 9A, JANUARY 1986

SHOWING MEDIAN AND 90th, 75th, 25th AND 10th PERCENTILE LEVELS

Student Populations - Board (extreme left) and 16 schools

TABLE 3

OVERVIEW OF DIFFERENCES IN MALE AND FEMALE MEAN SCORES
SYSTEM-WIDE STUDENT EVALUATION RESULTS, 1985

SUBJECT	ESTIMATED		ACTUAL		END OF YEAR	
	Deg. of Diff.	Higher Mean	Deg. of Diff.	Higher Mean	Deg. of Diff.	Higher Mean
Chemistry 12A	n.s.	-	**	M	n.s.	-
English 9A	***	F	n.s.	-	***	F
English 9G	n.s.	-	n.s.	-	n.s.	-
English 12A	***	F	*	F	***	F
Français 9A-1	***	F	***	F	***	F
Français 9A-2	***	F	***	F	-	-
Français 9G-1	**	F	**	F	***	F
Français 9G-2	***	F	n.s.	-	-	-
Français 12A	***	F	n.s.	-	***	-
Mathematics 9A	**	F	n.s.	-	**	F
Mathematics 9G-1	*	F	n.s.	-	n.s.	-
Mathematics 9G-2	n.s.	-	n.s.	-	-	-
Mathematics 12A	n.s.	-	n.s.	-	n.s.	-
Science 9A	**	F	n.s.	-	***	F

n.s. Non significant difference between means for males and females
* Significant difference between means for males and females $p < 0.05$
** Significant difference between means for males and females $p < 0.01$
*** Significant difference between means for males and females $p < 0.001$

for English and smaller for the sciences and Mathematics. The correlation coefficients based on general level scores are smaller than those based on advanced level.

TABLE 4

CORRELATION COEFFICIENTS FOR FRANCAIS 12A
BOARD-WIDE EXAMINATION RESULTS, 1985

School	Actual and Estimated r	N	Actual and End-of-Year r	N	Estimated and End-of-Year r	N
022	0.77	123	0.84	118	0.87	117
031	0.74	99	0.82	98	0.93	98
032	0.81	61	0.86	61	0.80	62
054	0.73	62	0.84	63	0.88	62

Test Specific Results. Depending on the nature of the test and the type of scoring used there may be other types of data produced. As an example, the test of Physics 11A given in June, 1986 included a student questionnaire and optically scanned multiple-choice items.

The questionnaire that made up the first part of the Physics test did not influence the test score. It was composed of 22 multiple-choice items and was designed to collect information about students' course selections, study times, career choices and attitudes towards

physics. Since the 1985-86 enrolment of female students in this subject was about 65 per cent that of male students in grade 11 and about 54 per cent in grade 13 there was considerable interest in the responses of each sex. The percentage counts for four items on the questionnaire are shown in Table 5. These answers suggest that the female students do not like the course quite as well as the male students do although this is not because they feel the course has been designed for male students. The females study longer hours and feel more strongly than the males that females are as able as males to learn physics.

When an optical scanner scores some or all of a test there is a printed listing produced automatically for each school with individual student responses for each item as well as student total scores, class means and item response data. This information allows examination of the test results in greater depth at either the individual, class or board level. For instance, it is possible to calculate the mean score for the board and for each school on groups of items representing a particular topic or skill. Table 6 shows such means broken down further by sex for the results of one part of the Physics test. Table 7 gives the means by sex for the entire test of Physics.

Rescoring. During the summer following the first year of testing, a sample of Mathematics 12A and Chemistry 12A tests were rescored. Both tests had a large number of multiple-choice items which did not require rescoring so that one-half of the test by item value was rescored in Chemistry and 60 per cent in Mathematics. Three experienced teachers were assigned to rescore the Mathematics and two to Chemistry.

TABLE 5
RESULTS OF STUDENT QUESTIONNAIRE IN PHYSICS 11A
SWSE TEST OF JUNE 1986

Questions and Responses	Percentage of Responses		
	Males (N=552)	Females (N=440)	Total (N=1053)
A. How well do you like the physics course you are taking?			
a. Very well	20	16	18
b. Reasonably well	46	40	43
c. Not too well	13	17	15
d. Not at all	4	5	4
e. This varies greatly with different parts of the course	15	20	18
No response	2	2	2
B. What is the usual amount of time you spend each week in studying or preparing for physics outside of the classroom?			
a. Less than one hour	34	19	27
b. Between one and two hours	37	37	37
c. Between two and three hours	19	26	22
d. Between three and five hours	6	14	9
e. More than five hours	3	3	3
No response	2	1	2
C. The physics course you are now taking is designed to appeal equally well to students of either sex.			
a. Strongly agree	58	70	64
b. Mildly agree	22	18	20
c. Uncertrain	12	8	10
d. Mildly disagree	3	1	2
e. Strongly disagree	1	0	1
No response	3	2	3
D. In general there is no difference between the abilities of males and females to learn physics.			
a. Strongly agree	63	78	69
b. Mildly agree	16	11	14
c. Uncertain	13	5	9
d. Mildly disagree	5	2	3
e. Strongly disagree	2	2	2
No response	2	2	3

TABLE 6

RESULTS BY TOPIC AND SEX
FOR MULTIPLE-CHOICE ITEMS IN PHYSICS 11A
SWSE TEST OF JUNE 1986

Topic	Item Number	Percentage Correct	
		Males	Females
Mechanics		67.7	65.8
a. Measurement, Significant figures & Units	41, 42, 52, 54, 81	(74.0)	(76.6)
b. Graphics	43, 44, 51	(59.0)	(55.3)
c. Others	45, 46, 47, 48 49, 50, 53, 55 56, 82, 83, 84 85, 86, 87	(66.0)	(64.3)
Waves and Light	57, 58, 59, 60 61, 62, 63, 64 65, 66, 88, 89 90, 91	66.4	63.6
Electricity	67, 68, 69, 70, 71, 72, 92*	54.1	56.9

* 93 & 94 are not included in this table.

TABLE 7
RESULTS BY SEX FOR PHYSICS 11A
SWSE TEST OF JUNE 1986

	MEAN	SD	N
Males	61.2	17.1	584
Females	60.6*	16.5	470

* The difference between the means is not significant.

Although most papers in the sample were scored once there was an overlap of between 40 to 50 papers for each pair of scorers. The sample size from each school was about 40 unless fewer students from that school wrote the test.

The results of the rescoring showed the correlation between the original marks and the rescorers' marks for each school to be at least 0.95 (in a possible range of zero to one). This large correlation coefficient indicates that the ranking of the sampled students in the school changed very little after the tests were rescored.

When the mean for the sample from a particular school was compared to the corresponding mean of the original scores for the same group, the differences varied somewhat from school to school. The largest difference on this basis was four points and the smallest was 0. That means that in the most extreme case, the average difference for a school's sample of the teacher-determined marks and the rescorer-determined marks was 4 points in 100. In general, the differences were smaller in Mathematics than in Chemistry.

The results of this rescoring experience satisfied those concerned that there was not a large diversity in scoring practices and there has been no subsequent rescoring conducted.

PROFESSIONAL DEVELOPMENT ACTIVITIES

One important aspect of the SWSE project is the opportunity it provides for teachers and others to take part in shared activities that are designed to promote skills in evaluation. Such activities have been on-going since the project began and the nature of some of these will be described in the following paragraphs.

Fall Conference. In September 1985, the beginning of the second year of the project, a two and a half day conference was held for members of the test construction committees. A specialist in educational measurement and evaluation from a nearby university, planned the agenda and was the major speaker. He spoke on using evaluation data to improve programmes and then directed working sessions for which participants were grouped according to their subject area.

There was a very positive response to this conference and the organizer has returned to the OBE for several shorter meetings related to the project.

Test Construction Committee. Participation on a test construction committee is usually viewed by teachers as being a very valuable learning experience. They welcome the opportunity to talk in depth with colleagues about the course objectives and methods of evaluation and emerge from the experience feeling their professional skills have been strengthened. Most find they have put extra hours into the committee work but feel they would make the same choice again.

Teacher Workshops. The SWSE project has been the impetus and source of funds for many teacher workshops. They may be held for the staff of a single school, for heads of departments or for all teachers of a subject. These are generally planned by consultants and often focus on some aspects of the evaluation process.

Frequently a meeting of teachers takes place within a day of the SWSE test to review the prescribed marking scheme and clarify any uncertainties or controversies. When such meetings follow the test, they are also an opportunity for teachers to react to the test and serve to inform the next test construction committee about features of the test that were either liked or disliked.

In the language arts, workshops are often held well in advance of the tests to prepare teachers for more subjective marking. Sometimes presentations are made by specialists from outside the board. Samples of student writing are often distributed so that all present can react to the same piece. Through discussion and sharing views on many different types of students' work, teachers gradually move towards a general consensus on scoring criteria and standards.

During the second year of the project, half of the SWSE test mark in English was based on the writing folder. This contains work that the student has written over the year and may include outlines, drafts and final copies. The method of producing a score for the writing folder became the topic for several workshops for English teachers.

Another type of workshop was held at the beginning of the third year for about 50 Mathematics teachers who met for a half-day and then had dinner together. They heard a short talk by an outside speaker and

then broke into groups of about six or seven. Each group was asked to study the objectives for a particular section of the Mathematics 12A course and then develop items to test the objectives. The items produced by each group were then presented to all present and sometimes modified through suggestions from the floor. The products of this workshop will be available to the test construction committee.

Pooled marking in English. During the first year of the SWSE project, the test committee for English 12A decided to have an essay on the test marked anonymously by two teachers other than the student's own teacher. The essay was worth one-half the total value of the test and the students had the choice of ten possible topics.

A procedure was set up for this process which included having the student's number rather than name and school appearing on the essay booklet so that the essay would be marked anonymously. The papers were sorted into bundles of about 35, with a mixture of schools represented in each, and then each bundle was delivered to two different teachers for marking. The teachers were instructed to assign to each composition a holistically determined mark (that is, based on an initial impression) falling between 1 and 10 inclusive. They had attended a workshop on holistic scoring before the scoring began.

Other stages in the process included refereeing, or resolving a difference of 4 or more points in the two teachers' scores, and callibration, or determining if the quality of writing at each point on the scale was the same for all the topics. In the majority of cases the average of the two scores became the students mark and for most topics no callibration was needed.

Due to the exploratory nature of this undertaking and the fairly large number of over 1300 essays to be processed the operation became somewhat cumbersome. The process was completed, however, and provided teachers with an outside evaluation of their own students' work as well as a chance to read a selection of writings from other schools.

Post-test meetings. At some time following most SWSE tests there is a meeting with teachers at which the results are discussed. The type and extent of the follow-up vary greatly and depend somewhat on the type of test.

Much more information about student responses to individual items is available on multiple-choice items scored by an optical scanner. When a follow-up meeting includes the distribution of the output listings for each school there is usually a member of the Research Department present to go over the material and suggest possible uses for it. Such meetings were held for the Chemistry, Physics and Geography tests.

With the results of the optical scanning, teachers have a record of the response given by each of their students to every item in this section. They also have a count for each class and the whole school for every possible response. With this information they can analyze the results in many ways such as grouping items by topics as shown in Table 6 on page 25. Such a table with the board mean given and a blank space for the school mean, is typically distributed for teachers to complete for their own school at post-test meetings.

The accumulated optical scan data from all schools may be used in an assessment of individual test items based on the discrimination

index and selection of alternative answers. This leads teachers to examine the strengths and weaknesses of the items, suggest improvements and thereby increase their own expertise in constructing multiple-choice questions.

The rescoring of Chemistry tests led to an interesting presentation at a meeting of Chemistry teachers by one of the rescoring team. This teacher had noted a great variation in the instructive value of the teachers' marks on the papers and wished to stress the importance and illustrate methods of using scoring as a teaching device. Part of the presentation was a review of the rescorer's decisions in controversial marking situations. The other teachers' reactions to some decisions showed that this group still had differences in their scoring practices and beliefs. Many teachers commented at the end of the meeting, however, that they had found the afternoon to be extremely valuable and that they would welcome other meetings of this type.

Because the mean of one of the tests in Français was thought to be unusually low, a member of the Research Department was asked to prepare a guide for examining the test in order to determine possible causes for the poor results. This guide was used by the consultant at a follow-up meeting with the teachers.

In-school Meetings. The principal of a school receives the test results and then meets with teachers and department heads to discuss them. If the results are not satisfactory, changes are considered. These may be in the balance or extent of the course content, the evaluative procedures, teaching strategies, resources or other areas.

Consultants go to the schools to meet with teachers and others for discussions of the test results and appropriate responses. The test results are also examined in meetings between the superintendent for the school area and the principal, with or without school teaching staff.

At times information arising out of the SWSE project is sent in a written form to the school. For instance, one result of rescoring the tests in Mathematics was to prepare a description of proper form for presenting calculations and solutions. This was the basis of small group discussions in some schools.

VIEWS OF PARTICIPANTS

The SWSE project has been very wide in scope and has involved hundreds of professional educators as well as thousands of students in 21 high schools of the board. In this final section of the report an attempt will be made to present some views of the project from various perspectives. The opinions and attitudes have been collected through working as part of the project team and are not based on results of a formal survey.

The Administration. The administration has seen the aim of this project to be in promoting a high level of academic standard in all schools through establishing more uniform evaluative procedures. Using system-wide tests in certain core subjects was one important feature of the project but another integral part of the process was developing teachers' skills in evaluation.

The project was not initiated by the administration although they did prepare the plan which was finally accepted. This was smaller in scope than the testing program the trustees had originally proposed. The administration realized that the imposition of external testing could cause a great deal of resentment and anxiety in the teaching body and tried from the beginning to implement the project in such a way that this would be minimized.

The majority of committee members are classroom teachers whose ideas and suggestions guide the development of the instrument. Feedback to other teachers of information related to the test is part of the process. Teachers generally mark the tests of their own students and always submit the students' scores themselves. There has been support for most teacher activities related to the testing and the form of these is influenced by teachers' suggestions. The results of the tests are distributed so that information is given only to the persons directly involved. Efforts have been made to avoid comparisons of schools and any association of teachers with test scores. The results of this approach and increasing familiarity with the project has resulted in a higher degree of acceptance by teachers than was seen initially.

In guiding the project over the first months there also seems to have been a shift of emphasis on the part of the administration. There appeared to be more concern about the testing component of the undertaking at the beginning but after two years more value seems to be placed in promoting evaluative expertise in the schools. Teachers are meeting to consider course objectives, teaching methods and evaluation techniques and the benefits of these exchanges should have an effect in the classroom. The SWSE tests are still the impetus for the activity

but there is somewhat less concern for the scores themselves as well as the slight variations in applying the marking criteria that may occur from one school to another. Because of the spin-offs from the project, such as installing common course outlines and upgrading teachers' evaluative skills, the administration appears to be more favourably disposed to SWSE now than at the beginning and has recommended its continuation for three more years.

Consultants. Consultants have played a leading role in carrying out the mandate of the SWSE. They are positioned between the teachers and the project planners which usually gives them a broad perspective, particularly over their own subject area. Most consultants are aware of the advantages and disadvantages of the project to the various people they contact and serve. Although they are all supportive in carrying out the project operations they would probably be divided in recommending its continuation.

Certainly the consultants use a considerable amount of time in meeting the requirements of the SWSE. On the other hand the project has provided the consultants with some additional opportunities and funds to use in reaching their normal goals which are independent of the project. For courses being tested by SWSE, the consultant's task may be eased because teachers have more reason to study objectives and adher to course outlines. Due to SWSE results, consultants have supplementary information about certain aspects of the schools they serve which could be of value in their work.

Principals. Principals were very apprehensive about the SWSE project before it began but this seems to be somewhat reduced. They were quite

naturally worried about the use that could be made of the results and the possibility of inaccurate conclusions being drawn, especially in trying to rate schools. They were also concerned about the effects of the tests on students and on teachers. They feared the additional testing would be very disruptive and, in general, preferred the SWSE testing to be conducted in June. Many principals are in agreement with teachers who state that the usual type of external testing is not appropriate to general level students.

Teachers. When the SWSE project was introduced, very few teachers were in support of it. They feared that an external test would impose an excessive amount of control over their classroom activities which would become "test-driven". Teachers felt they might lose the opportunity to respond to special circumstances. For instance, many Physics teachers claim to have been very frustrated as they were preparing classes for the grade 11 SWSE test just as the nuclear accident at Chernobyl occurred.

Teachers were also concerned about the potential of using such tests to evaluate professional skills. At a time of declining enrolment, school closures and reorganization of local school boards there was already a great deal of insecurity and it appeared that testing might pose an added threat. Teachers also anticipated increased pressure from superiors to have students perform well on SWSE tests.

There was the fear that results would used in a way that is unfair to students, teachers or schools. The schools in the board vary widely and may draw on quite different populations. Some schools have enriched programs while others specialize in commercial or technical training. The socio-economic characteristics of the student bodies

differ and there are both French and English high schools. A few teachers welcomed the opportunity to compare their own students with board means but most did not feel this information would compensate for the disadvantages of the program.

As the project has progressed there has been a change in the attitudes of teachers, particularly in some subjects. Many of those on test construction committees have felt they have gained considerable expertise in evaluation as a result. Their generally positive attitudes towards the project have reassured some of their colleagues. Other teachers have felt benefits from some of the SWSE-related activities.

Particularly in the case of teachers of English has there been a very extensive shift in feelings towards the project. This group was particularly opposed to a system-wide test at the beginning but, by the end of the second year, the strong opposition had almost completely disappeared. Some of the factors that caused this change seem to have been the numerous workshops and meetings required to support the testing process, the acceptance of the project by committee members, the appreciation of the quality of the tests and the realization that teachers were still the evaluators.

The tests themselves have usually been well accepted by teachers. Generally they are seen to be fair and well constructed. In some cases the teachers are pleased to have the test as a model or as a source of items for future tests. There has been some controversy over the use of multiple-choice items. They had some limited use in English but this has been discontinued. Also the teachers of Mathematics 12A asked that they not be used in the next test. It was reported, however, that

at least some of the students writing the Geography 9G test were very pleased with the multiple-choice format and particularly with the maps and charts which accompanied the items.

One of the objections of teachers to the SWSE has been due to the choice of courses to be tested. They feel that a course that is soon to be discontinued or one that has just been installed should not be involved in an external test. Teachers argue that resources are wasted on outgoing courses and that it would be more beneficial to prepare course outlines and instruments for an incoming course. They note that it often takes several years to actually complete a course changeover. In the sciences, for instance, new equipment must be obtained and new experiments set up. In other subjects too, it may not be possible to purchase all the new textbooks and classroom materials in the same year. In addition, teachers feel that there are often increased pressures involved in introducing a course for the first time and that it becomes increasingly stressful if the classes must also prepare for an external test.

Some teachers are concerned about the proportion of the student's final mark that will be determined by the SWSE test. This is especially true at the higher grade levels where the student's future may be at stake. For this reason there has been discussion about allowing teachers of grade 13 courses to see the SWSE test beforehand.

Students. For the most part there has not been a strong reaction by students to the project. However, there is greater apprehension about the SWSE tests than there would be for tests constructed within the school for the same subject. Also, some students had formerly been

excused from writing tests in June if their course mark was above a certain level and now they have lost this privilege for SWSE tested courses.

On the other hand, the greater degree of uniformity in courses due to establishing common course outlines makes it less disruptive for a student transfering between schools. It is also reassuring for some students to know that the mark they receive has the same value no matter which school they attend.

Students did protest to their teachers against having a questionnaire included on the Physics tests. There were also a few essays written by students which reflected their negative view towards board-wide testing. In general however, the students seem to have taken the testing in their stride.

CONCLUSION

This report has described some aspects of a board-wide testing program from the perspective of one of the many persons involved in it. The intention has been to give an outline of the procedures followed, a sample of the types of results obtained and an impression of the reactions of the project participants.

One of the notable features of the SWSE project is the degree of flexibility that it has incorporated. The many committees have been allowed to work relatively independently, each choosing procedures and materials felt to be most suitable for their own particular subject and situation. This approach appears to have served the overall purpose of promoting excellence in evaluation very well. The flexibility has also allowed the project to gradually evolve in response to on-going experi-

ences and an increasing number of possibilities for meeting needs and objectives. In the area of follow-up activities alone there will undoubtedly be many new developments in the next few years. Like many educational endeavours, however, the most important feature is the involvement of teachers which has been essential in the project's operation to date and will no doubt determine its ultimate usefullness in the promotion of high academic standards in the schools.

APPENDIX

TABLE A1

OUTLINE OF TESTS USED IN THE SWSE PROJECT, 1984-85 AND 1985-86
ADVANCED LEVEL COURSES

Part	Description of item(s)	Marking	Value	Time (minutes)
English 9A, April 25 and May 2, 1985				
A	Sight passage of prose with 10 multiple-choice items	2 points each	20	30 (Apr)
B	Composition: Rough draft - April Final copy - May	Scheme provided with breakdown for content organization and style	30 50	60 (Apr) 60 (May) 90 (Apr) 60 (May)
English 11A, June 5, 1986				
A	Essay based on Shakespeare or a novel.	Scheme provided for each of 4 alternative questions.	30	60
B	Sight poem with 10 multiple-choice items	1 point each	10	20
C	Short essay based on a sight prose passage	Scheme based on construction, content and technical aspects	10	40
D	Writing folder essays		50	not in exam time
			100	120
English 12A, March 5, 1985				
A	Argumentative Essay (10 topics to choose from)	Pooled marking Holistic method	50	45
B	Sight poem with 10 multiple-choice items	2 points each	20	25
C	Essay on a novel or a Shakespearian play	Marking scheme provided based on introduction, development, references to text and writing style	30	50
			100	120

TABLE A1 (continued)

OUTLINE OF TESTS USED IN THE SWSE PROJECT, 1984-85 AND 1985-86
ADVANCED LEVEL COURSES

Part	Description of item(s)	Marking	Value	Time (minutes)
English 12A, June 5, 1986				
A	Essay based on Shakespeare or a novel	Descriptive guide and a breakdown for content and organization/mechanics	30	60
B	Sight poem with 10 multiple-choice items	1 point each	10	20
C	Short essay based on a sight prose passage	Scheme with breakdown for construction, content and technical aspects	10	40
			50	120
Français 9A, January 16, 1985				
A	Listen to recorded instructions, mark a map as instructed	10 specific instructions to check	20	5
B	Sight reading of an article	Short answer items with mark breakdown for each	30	15
C	Grammar	Fill in blanks with verbs and homonyms	20	10
D	Write a letter on 1 of 2 topics (3 paragraphs minimum and about 200 words)	Scheme with breakdown for letter format, variation of sentences, construction, content and grammar	30	20
			100	50 (+10)
Français 9A, May 2, 1985				
A	Sight reading of prose	Short answer items with a detailed scheme	50	40
B	Grammar - homonyms, agreement, verbs and constructing compound sentences	Homonyms - 0.5 pt Correct words - 1 pt Correct sentences - 1.5 pts	40	15
C	Write a dialogue-choice of two situations	Scheme based on content, setting, syntax, punctuation and grammar	60	35
			150	90

TABLE A1 (continued)

OUTLINE OF TESTS USED IN THE SWSE PROJECT, 1984-85 AND 1985-86
ADVANCED LEVEL COURSES

Part	Description of item(s)	Marking	Value	Time (minutes)
Français 11A, January 15, 1986				
A	Sight reading of an article	Scheme for short-answer items and a short essay	120 (48%)	60
B	Grammar	Homonymes (15) Verbs (20)	35 (14%)	15
C	Composition - Continuing a given narrative text for 35 lines	Scheme provided with breakdown for development, style, sentences structure, vocabulary and grammar	95 (38%)	45
			250 (100%)	120
Français 12A, March 5, 1985				
A	Sight reading of a poem	Scheme for short-answer questions	20	20
B	Sight reading of an editorial	Scheme for short-answer questions	30	30
C	Opinion essay - Prepare a detailed outline and rough copy	Scheme broken down for outline, content, construction, style, syntax, and vocabulary	50	60
			100	110 (+10)
Français 12A, June 5, 1986				
A	Sight reading of a poem	Scheme for 8 short-answer items	15	20
	Sight reading of an article	Scheme for 10 short-answer items	25	30
	Mechanics of writing in Section A responses	0.5 marks deducted for each error	10	
B	Essay of 325 to 350 words	Scheme provided with breakdown for outline, development, style, sentences structure, vocabulary and grammar	50	60
			100	110 (+10)

TABLE A1 (continued)

OUTLINE OF TESTS USED IN THE SWSE PROJECT, 1984-85 AND 1985-86
ADVANCED LEVEL COURSES

Part	Description of item(s)	Marking	Value	Time (minutes)
Mathematics 9A, April 2, 1985				
A	Multiple-choice	1 point each	25	
B	Short-answer	Scheme showing requirements for full and partial marks	25	
C	Questions requiring complete solutions	Scheme giving requirements, in proper form and with all work shown, for full and partial marks	50	
			100	120
Mathematics 12A, March 1, 1985				
A	Multiple-choice	1 point each	12	
B	Short-answer (13 items, some with parts)	Scheme showing requirements for full and partial marks	28	
C	Questions requiring complete solutions	Scheme giving requirements, in proper form and with all work shown, for full and partial marks	60	
			100	120
Mathematics 12A, June 4, 1986				
A	Multiple-choice	1 point each	15	
B	Short-answer	Scheme showing requirements for full and partial marks	20	
C	Questions requiring complete solutions	Scheme giving requirements, in proper form and with all work shown, for full and partial marks	65	
			100	150
Calculus 13A, June 9, 1986				
A	Short-answer	No work shown 1 point each	10	
B	Short-answer	Work must be shown 2 points each	20	
C	Longer questions requiring complete solutions	Sheme showing work necessary to get full and partial marks	90	
			120	150

TABLE A1 (continued)

OUTLINE OF TESTS USED IN THE SWSE PROJECT, 1984-85 AND 1985-86
ADVANCED LEVEL COURSES

Part	Description of item(s)	Marking	Value	Time (minutes)
Science 9A, April 10, 1985				
A	Multiple-choice (40 items)	1 mark each - marked by teacher then optical scanner	40	40
B	Short-answer some with diagrams	Sample answers given with instructions for marking. Complete sentences required	60	70
			100	110 (+10)
Physics 11A, June 9, 1986				
A	Student questionnaire (22 items)	Not included in scoring	0	(15)
B	Multiple-choice (32 items)	1 point each Optical scanner	32	35
C	Multiple-choice (14 items)	2 points each Optical scanner	28	20
D	Short-answer and numerical problems	Scheme with requirements, including units, for full and partial marks	40	65
			100	120
Chemistry 12A, March 7, 1985				
A	Multiple-choice (30 items)	1 point each Optical scanner	30	40
B	Multiple-choice (10 items)	2 points each Optical scanner	20	20
C	Short answer	No half marks given	23	20
D	Numerical problems	Show calculations and units	20	25
			93	105 (+15)
Geography 9A, Janaury 15, 1985				
	Multiple-choice (100 items - many with maps and diagrams)	1 point each	100	90

TABLE A2

OUTLINE OF TESTS USED IN THE SWSE PROJECT, 1984-85 AND 1985-86
GENERAL LEVEL COURSES

Part	Description of item(s)	Marking	Value	Time (minutes)
English 9G, April 25, 1985				
A	Writing sentences describing a sketch	Based on correctness of observation and sentence structure	5	10
B	Write a paragraph on a given topic	Scheme based on content, style and organization	10	10
C	Completing a story in 2 paragraphs. Choose 1 of 3 beginnings	Scheme based on content, style and organization	15	15
D	Sight passage	Six short-answer and multiple-choice items	10	10
E	Sight poem	Six short-answer and multiple-choice items	10	10
			50	55 (+5)
English 9G, May 2, 1985				
The format was exactly the same as the English 9G test given one week earlier. The content was different.				
Français 9G, January 16, 1985				
A	Listen to recorded instructions	Fill in a chart -20 responses	10	5
B	Grammar - homonyms and verbs in specified tenses	Homonyms (10 x 1) Verbs (20 x 0.5)	20	10
C	Sight reading of an article	Short answer items (including completing a grid)	30	15
D	Write a letter to a friend on a given subject	Scheme based on the quality and quantity of information, introduction and conclusion, use of paragraphs and complete sentences and grammar	40	20
			100	50 (+10)

TABLE A2 (continued)

OUTLINE OF TESTS USED IN THE SWSE PROJECT, 1984-85 AND 1985-86
GENERAL LEVEL COURSES

Part	Description of item(s)	Marking	Value	Time (minutes)
Français 9G, May 2, 1985				
A	Listen to a recording (with expressive speech)	Short answer questions (20 responses)	10	5
B	Grammar	Homonyms (10 x 0.5) Word endings (10 x 0.5) Verbs (10 x 0.5) Compound Sentences (10 x 1)	25	10
C	Sight reading of a narrative text	Short answer questions on comprehension	25	15
D	Descriptive text on a given subject	Introduction-Conclusion (5) Structure of paragraphs (5) Complete sentences (5) Agreement, verbs, adjectives (5) Information given (20)	40	20
			100	50 (+10)
Mathematics 9G, January 23, 1985				
A	Objective items	Fill in blank (17) True-False (3)	20	
B	Questions requiring complete solutions	Scheme indicating requirements for full and partial marks	60	
			80	60
Mathematics 9G, April 2, 1985				
A	Statistics and Graphs (6 items, many with parts)	Scheme indicating requirements for full and partial marks	30	
B	Measuremnt and Geometry (Geometry sets required)	Scheme describing answers and drawings needed for full and partial marks	30	
C	Squaring and Square Roots (5 items, some with parts)	Scheme showing requirements for each point	20	
			80	60
Geography 9G, January 1986				
	Multiple-choice - 70 items (many based on maps, diagrams or tables)	1 point each	70	60
			70	60